

Genetic characterization of chestnut (*Castanea sativa* Mill.) orchards and traditional nut varieties in El Bierzo, a glacial refuge and major cultivation site in northwestern Spain

Julia Quintana · Angela Contreras · Irene Merino ·
Alvaro Vinuesa · Gabriela Orozco · Felipe Ovalle ·
Luis Gomez

Abstract This paper presents a detailed genetic study of *Castanea sativa* in El Bierzo, a major nut production region with interesting features. It is located within a glacial refuge at one extreme of the distribution area (northwest Spain); it has a centenary tradition of chestnut management; and more importantly, it shows an unusual degree of genetic isolation. Seven nuclear microsatellite markers were selected to analyze the genetic variability and structure of 169 local trees grafted for nut production. We analyzed in the same manner 62 local nuts. The selected loci were highly discriminant for the genotypes studied, giving a combined probability of identity of 6.1×10^{-6} . An unprecedented density of trees was sampled for this project over the entire region, and nuts were collected representing 18 cultivars marketed by local producers. Several instances of misclassification by local growers were detected. Fixation index estimates and analysis of molecular variance (AMOVA) data are supportive of an unexpectedly

high level of genetic differentiation in El Bierzo, larger than that estimated in a previous study with broader geographical scope but based on limited local sampling (Pereira-Lorenzo et al., Tree Genet Genomes 6: 701–715, 2010a). Likewise, we have determined that clonality due to grafting had been previously overestimated. In line with these observations, no significant spatial structure was found using both a model-based Bayesian procedure and Mantel's tests. Taken together, our results evidence the need for more fine-scale genetic studies if conservation strategies are to be efficiently improved.

Keywords Cultivar identification · Genetic structure · Simple sequence repeat (SSR) · Sweet chestnut · Human management

Introduction

It is difficult to overstate the ecological importance of Fagaceae trees as they dominate the hardwood forests of temperate regions. In addition, several members of this family have important economic uses (Kremer et al. 2012; Petit et al. 2013). The sweet chestnut (*Castanea sativa* Mill.) constitutes an outstanding example. Owing to its unique combination of highly valued nuts and timber, colonizing aptitude, and rapid growth, it has been actively dispersed across Europe since ancient times, especially during the Roman Empire (Conedera et al. 2004; Krebs et al. 2004). Intensive grafting and other agronomic practices have increased its importance ever since, fostering new plantations and the development of cultivars with improved nut quality, stress tolerance, and wood properties. The effects of such management remain largely unexplored at the molecular level (Mattioni et al. 2008), despite its likely influence on the genetic makeup of chestnut populations. The same is true for most long-lived woody plants

(Neale and Kremer 2011), although a few studies have targeted economically valuable species. Their conclusions ranged from a substantial human influence on the genetic structure of walnut (*Juglans regia* L.), irrespective of physical barriers (Pollegioni et al. 2011), to much less significant effects in cork oak (*Quercus suber* L.) or Mediterranean cypress (*Cupressus sempervirens* L.), which appear to be more reliant on natural history than management (Magri et al. 2007 and Bagnoli et al. 2009, respectively). Like the sweet chestnut, all these species are essentially wind pollinated. In the case of *C. sativa*, Mattioni et al. (2008) have pondered human influence across Europe, concluding that there is a strong geographical correspondence between inferred population structure and glacial refugia. The homogeneous gene pool observed by these authors in Italy and Spain, however, also reflected the impact of human-mediated colonization. Although this work did not include samples from the region studied here, a glacial refuge itself, expecting a mixed influence seems a reasonable hypothesis.

Recent studies using molecular markers have confirmed that the genetic landscape of *C. sativa* is complex across its distribution area—from the Atlantic Ocean to the Caspian Sea—especially in traditional nut production areas, where both grafted material and wild populations have coexisted for long periods of time (Gobbin et al. 2007; Lang et al. 2007; Mattioni et al. 2008, 2013; Martin et al. 2010). The northwest region of the Iberian Peninsula is particularly interesting in this context because (i) it is located at one extreme of the distribution area and concentrates the most extensive sweet chestnut stands in Spain and Portugal, (ii) it has been one of the major focal points for the native species after the last glaciation, and (iii) it has a centenary tradition of chestnut cultivation and management (Conedera et al. 2004; Ramos-Cabrer and Pereira-Lorenzo 2005; Pereira-Lorenzo et al. 2010a, 2011; Martin et al. 2012; Mattioni et al. 2013). This region hosts nowadays prime orchards for commercial nut production, located mainly in Galicia and El Bierzo (Castile and Leon) (Rubio 2009). Nuts with these provenances are widely valued in Spain, France, and Portugal for their organoleptic properties (Pereira-Lorenzo et al. 2006b), representing an important economic asset. Besides flavor and cooking properties, chestnuts are increasingly appreciated worldwide for their low fat and high fiber content. Moreover, they represent a rich carbohydrate source that does not upset people affected by celiac disease. According to FAO, the global gross production in constant terms has doubled in less than a decade (see <http://faostat.fao.org> for the latest data).

We report here a detailed genetic characterization of *C. sativa* in El Bierzo, a region where this species has been

much less studied than in Galicia (see Martin et al. 2012 and references therein). Located into a tectonic depression surrounded by mountains (up to 2100 m above sea level), El Bierzo provides excellent climatic conditions for chestnut growth and high-quality nut production (Rubio 2009; Pereira-Lorenzo et al. 2010b). More importantly, it also provides a genetically isolated scenario to study the effects of long-term management. Such isolation is reflected in the variety of local agricultural products officially protected with a designation of origin or a geographical indication, such as pears, apples, peppers, or chestnuts themselves (www.castanadelbierzo.es). Throughout this study, we used nuclear microsatellite markers specifically developed for *C. sativa* by Buck et al. (2003) and Marinoni et al. (2003). Also known as simple sequence repeats (SSRs), these polymorphic markers have become widely used to study population genetic diversity, parental relationships, and patterns of gene flow, among other applications (Ellegren 2004). Whereas their development is still limited in forest trees (Neale and Kremer 2011), microsatellite loci have been successfully used in recent years to improve our overall understanding of the genetic landscape of sweet chestnut in Europe (e.g., Mattioni et al. 2008; Martin et al. 2010; and others). These markers have also been applied to genotype germplasm or differentiate chestnut cultivars, often at national or regional scales (e.g., Buck et al. 2003; Marinoni et al. 2003, 2013; Gobbin et al. 2007; Martin et al. 2009, 2012). Like in many other traditionally managed sites, cultivar identification in El Bierzo is still solely based on morphological and agronomic traits. Complementing this information with molecular marker data should largely facilitate the identification and cataloguing of local cultivars.

To date, some local *C. sativa* samples have been analyzed in El Bierzo using molecular markers, but only as part of broad studies aimed at establishing large-scale genetic structure and relationships (Pereira-Lorenzo et al. 2010a, 2011; Martin et al. 2012). Germplasm and nuts are thus poorly characterized in this region, which awaits more focused analyses with sufficiently large samples. The lack of fine-grained genetic data remains a major hurdle to establish a core collection of local varieties (Brown and Spillane 1999). Such collection will provide the fundamental data for optimal sampling, use, and conservation of local gene pools. It will also help improve the protection status for local nuts, currently based on a declaration of geographical provenance. Finally, a better genetic characterization will prevent the irreversible loss of local varieties harboring important adaptive or technological traits. The aim of this study was to use polymorphic microsatellite markers to analyze in depth the genetic structure and diversity of El Bierzo orchards (grafted trees for nut production). In addition, a representative sample of all traditional nut cultivars was analyzed with the same tools.

Material and methods

Plant material and field sampling

A total of 169 chestnut trees (*C. sativa* Mill.) were sampled in El Bierzo region (Castile and Leon, Spain). This is a major production site in Spain for high-quality nuts (Pereira-Lorenzo et al. 2010b), which are commercialized under the regulated geographical denomination *Castaña del Bierzo*. In contrast to intensively managed modern plantations, chestnut trees are exploited here under a traditional scheme, with numerous local farmers owning relatively few trees. All accessions analyzed came from grafted trees that have long been exploited for nut production (most were centenary), interspersed in many cases with wild chestnuts and other long-lived woody species such as *Quercus robur*, *Quercus pyrenaica*, *Alnus glutinosa*, or *Fraxinus angustifolia* (Rubio et al. 2009). We sampled all areas where *C. sativa* is the dominant species (some 18,900 ha), according to the latest National Forest Inventory (IFN3 1997–2007; official website: www.magrama.gob.es). Sampling was arranged in circular parcels so that one tree was selected at random and georeferenced every 2–3 ha, provided it was grafted and its diameter at breast height was at least 20 cm. As a result, the entire region was mapped with an unprecedented density, including all relevant orchards (Fig. 1). While tree age and cultural factors make graft sources untraceable, sampling distances were large enough to prevent an overrepresentation

of genetically homogeneous material. Five healthy leaves were harvested per selected tree between May and October of 2009 and 2010. Leaf material was placed in tagged plastic bags, transported to the laboratory on dry ice, and stored at -20°C . Furthermore, the main local growers provided a set of freshly harvested fruits (November 2011) belonging to the 18 traditional denominations that are locally commercialized: Courcla (3 fruits), De Presa (5), Galeguiña (3), Injerta (5), Marela (4), Marela Seroda (2), Mi Tierra (2), Negral (3), Parede (19), Raigona (3), Rapada (1), Rapiega (1), Rapona or Pelona (1), Roja (2), Seroda (1), Serodia (1), Valdecañada (3), and Verdella (3)—some of these denominations are also used in Galicia (Pereira-Lorenzo et al. 2006b; Pereira-Lorenzo et al. 2010b). Local production is rather limited for some of them, such as Rapada, Rapiega, Pelona, or Seroda, while Parede is widely cultivated and exported, mainly to France. Fruits were stored at -20°C until processed.

DNA extraction and SSR analysis

Prior to DNA extraction, chestnut leaves and cotyledons were ground to a fine powder in liquid nitrogen. Total genomic DNA was extracted from ~80 mg of powder with the Qiagen DNeasy Plant Mini Kit, according to the manufacturer's instructions (Qiagen). To select the most informative microsatellite loci for El Bierzo material, we evaluated the CsCAT and EMC primer pairs developed respectively by Marinoni et al. (2003) and Buck et al. (2003). The loci eventually selected for

Fig. 1 Geographical distribution of the chestnut trees analyzed in this study across El Bierzo (Castile and Leon, Spain). Local municipalities (black lines) and the position of each sampled tree (black triangles) are indicated. The tiles where the sweet chestnut is the dominant species are shaded in grey, according to the latest National Forest Inventory (IFN3), available at www.magrama.gob.es.



their polymorphism and consistency are presented in Table 1. A single locus was selected per linkage group. PCR amplifications were performed in a final volume of 10 µl containing 200 µM of each dNTP, 0.5 µM of each primer, 10 ng genomic DNA, 0.38 units *Taq* DNA polymerase (Biotools), and 1 µl of 10× reaction buffer (1× is 75 mM Tris-HCl, pH 9.0; 2 mM MgCl₂; 50 mM KCl; 20 mM ammonium sulfate). Amplification cycles consisted of an initial step of 5 min at 94 °C, followed by 30 cycles of 30 s at 94 °C, 30 s at the annealing temperature of each primer pair (Table 1), and 30 s at 72 °C, with a final extension step of 20 min at 72 °C. For multiplexing PCR, forward primers were labeled with 6-FAM, PET, NED, or VIC fluorescence dyes (Applied Biosystems). The amplified products were separated by capillary electrophoresis in an ABI 3730 Analyzer (Applied Biosystems), and fragment sizes were analyzed with the Peak Scanner 1.0 software from Applied Biosystems. Data quality was assessed with Micro-Checker 2.2.3 (Van Oosterhout et al. 2004) using 10,000 runs for the Monte Carlo simulations (95 % confidence level). To ensure that all loci were scored consistently, at least three independent replicates were performed for each amplification reaction.

Genetic diversity

The number of alleles per locus, the allelic frequencies, and the allelic richness were calculated with FSTAT 2.9.3.2 (Goudet 1995). Micro-Checker was used to calculate the frequency of null alleles using the heterozygote deficiency estimator of Chakraborty et al. (1992). PowerMarker 3.25 (Liu and Muse 2005) was used to determine the polymorphism information content (PIC; Guo and Elston 1999) of each locus. The ability of each marker to discriminate between two random genotypes (probability of identity, PI) was

calculated with GenAEx 6.5 (Peakall and Smouse 2006, 2012). The number of genotypes was calculated and used to assess clonality. Arlequin 3.5.1.3 (Excoffier and Lischer 2010) was used to estimate the observed (*H_o*) and expected (*H_e*) heterozygosity values. The inbreeding coefficient (*F_{IS}*) was computed with the locus-by-locus analysis of molecular variance (AMOVA) implemented in Arlequin (10⁵ permutations). Physical localization of loci was inferred from the consensus linkage map of Barreneche et al. (2004) and Casasoli et al. (2006).

Genetic structure

To identify the primary structure across the entire distribution, a Bayesian clustering analysis of individual genotypes was performed for the 169 leaf samples using STRUCTURE 2.3.4 (Pritchard et al. 2000; Hubisz et al. 2009). Ten independent runs for each *K* value (1–20) were completed to ensure reproducibility, using a burn-in of 100,000 followed by 10⁶ Monte Carlo Markov chain (MCMC) iterations. The admixture model with correlated allele frequencies was used without prior information on sampling location. The most likely *K* value was inferred using the Structure Harvester web server (Earl and von Holdt 2012), which implements the method described by Evanno et al. (2005). Pairwise distances based on fixation index (*F_{ST}*) values were calculated with Arlequin to evaluate between-cluster differences, with *p* values being estimated using 10⁵ permutations. Hierarchical structure was characterized using the AMOVA tool implemented in Arlequin. The significance of covariance components of the two levels of structure, inter-cluster and intra-cluster differences, was estimated with 10⁵ permutations. The covariance components were used to estimate overall *F_{ST}*. To determine whether the genotypes found for local fruits could be assigned to the above

Table 1 Summary statistics for the 7 SSR markers used to genotype 169 trees representing chestnut distribution in El Bierzo

Locus	LG	Ta	Sample size	<i>A</i>	<i>R</i>	<i>r</i>	PIC	PI	<i>N</i>	<i>H_o</i>	<i>H_e</i>	<i>F_{IS}</i>
CsCAT1	C8	52	169	4.99	3	−0.01	0.25	5.5E−01	7	0.28	0.27	−0.02
CsCAT2	C10	54	166	10.99	8	−0.06	0.65	7.7E−02	23	0.75	0.69	−0.09
CsCAT3	C12	54	169	9.95	6	0.07	0.54	1.7E−02	18	0.45	0.51	0.12*
CsCAT6	C1	50	168	8.00	3	−0.08	0.68	2.0E−03	21	0.85	0.73	−0.16
CsCAT8	C6	54	165	6.00	1	−0.06	0.65	2.8E−04	14	0.79	0.69	−0.13
CsCAT17	C2	54	168	7.00	2	−0.03	0.71	2.9E−05	18	0.80	0.75	−0.05
EMCs38	C4	56	168	11.96	7	0.24	0.54	6.1E−06	22	0.35	0.56	0.39**
Mean	—	—	—	8.41	—	—	0.57	—	17.57	0.61	0.60	0.06

CsCAT and EMC loci were developed, respectively, by Marinoni et al. (2003) and Buck et al. (2003). Linkage group assignment was based on Barreneche et al. (2004). Additional data are given in Online Resources 1 and 2

Ta annealing temperature, LG linkage group, *A* allelic richness (based on a minimum sample size of 165 individuals), *R* number of rare alleles (<5 %), *r* frequency of null alleles, PIC polymorphism information content, PI probability of identity for increasing locus combinations, *N* number of different genotypes, *H_o* observed heterozygosity, *H_e* expected heterozygosity, *F_{IS}* Wright's inbreeding coefficient

p* value <0.05, *p* value <0.01 in locus-by-locus AMOVA for deviation from Hardy-Weinberg proportions

clusters, a similar Bayesian approach was followed using the entire genotype pool. The STRUCTURE software was run as described above with no prior information. To examine if the resulting genetic structure was correlated with geographical distribution, Mantel's tests were performed with GenAlEx 6.5 (10^4 permutations). Moreover, a spatial Bayesian clustering algorithm was applied using BAPS 5.2 (Corander et al. 2008; Cheng et al. 2013) and performing 10 independent runs for each K value (1–20). The resulting partitions were visualized as Voronoi diagrams (tessellations).

Characterization of local fruits

To characterize the relationships among fruit cultivars, a study based on individual allelic frequencies was carried out, using the D_A distance (Nei et al. 1983) and the neighbor-joining algorithm (Saitou and Nei 1987). Bootstrap tests (Felsenstein 1985) were performed using 1000 replicates with the POPTREE2 software (Takezaki et al. 2010). Phylogenetic relationships were depicted as an unrooted tree and visualized with MEGA 5 (Tamura et al. 2011). Tree topology was used to classify the local varieties as well as to identify nomenclature problems. Furthermore, the average allele number as well as the observed and expected heterozygosity for each cluster was calculated with Arlequin. Within-group F_{IS} values and their significance were evaluated over 10^5 permutations. Bayesian clustering analyses were performed as described above, except that the traditional nomenclature was considered here.

Results

Genetic diversity analysis

Prior to multiplexing, a random set of 30 local trees grafted for nut production was analyzed to identify the most informative microsatellite loci. Among those developed by Marinoni et al.

(2003) and Buck et al. (2003), seven loci were selected for our large-scale study (Table 1). All loci are distributed over seven linkage groups (Barreneche et al. 2004). The variability of each locus was analyzed across the entire El Bierzo region, using leaf material harvested from a sample of 169 grafted trees (Fig. 1; see “Material and Methods” for sampling details). A summary of the genetic diversity found is presented in Table 1. The number of alleles per locus ranged from 8 to 12, with an average value of 8.4. The occurrence of null alleles was considered negligible after appropriate re-amplifications for loci CsCAT3 and EMCs38 under low stringency (Table 1). Allele frequencies ranged from 0.846 (CsCAT1) to 0.003 (several loci) (Online Resources 1 and 2). Of the 59 alleles found, 36 (61 %) were rare, i.e., showed a frequency below 0.05; 6 alleles (10.2 %) were genotype specific and only 3 alleles (5.1 %) showed frequencies above 0.60. In line with these figures, the average PIC, a common measure of the usefulness for linkage studies, was rather high for the entire marker set (0.57). CsCAT17 (PIC=0.71) and CsCAT1 (PIC=0.25) appeared to be the most and least informative loci, respectively. The probability of identity was 6.1×10^{-6} when the whole set of markers was considered.

The combination of profiles across all loci resulted in 98 different genotypes, 88 of which were represented by single individuals. The number of individuals sharing the remaining 10 genotypes was 2 (5 instances), 3, 4 (2 instances), 13, and 47. In agreement with the above PIC estimates, CsCAT17 showed a remarkable discriminative power (18 genotypes), while the opposite was true for CsCAT1 (only 7 genotypes). CsCAT2, EMCs38, CsCAT6, and CsCAT3 were also highly informative (23–18 genotypes).

The observed heterozygosity (H_o) values ranged from 0.278 (CsCAT1) to 0.798 (CsCAT17), with an average of 0.61. Expected heterozygosity (H_e) varied from 0.272 (CsCAT1) to 0.754 (CsCAT17), with an average of 0.60 (Table 1). EMCs38 showed the largest H_o - H_e difference among all markers analyzed. Calculations of F_{IS} , a measure of heterozygote deviation from Hardy-Weinberg equilibrium,

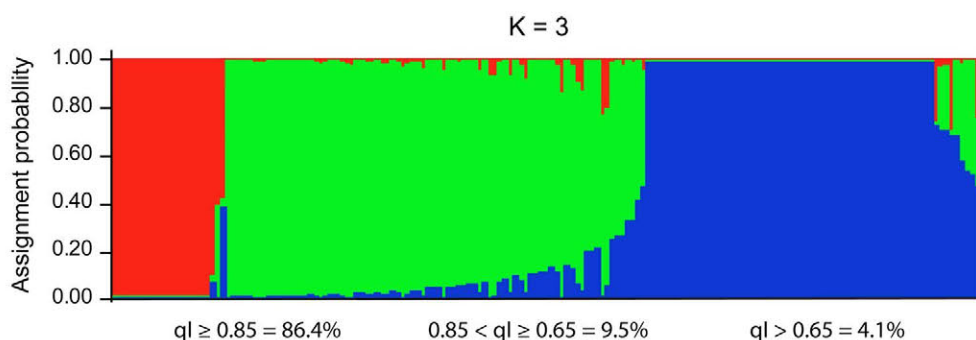


Fig. 2 Quantitative analysis of genetic structure according to the Bayesian clustering method implemented in STRUCTURE. The analysis of 169 chestnut samples was based on seven microsatellite markers. Each vertical bar represents a genotype, and bar colors show

the assignment probability to each of the three predicted clusters (see text). The fraction of genotypes with assignment probabilities equal to or higher than 0.85, between 0.85 and 0.65, and below 0.65 is expressed as percentage

rendered statistically significant deviations for CsCAT3 and EMCs38. The remaining loci showed heterozygote excess or no deviation. The average F_{IS} value was close to zero (Table 1). These figures are not unexpected for an outbreeding species long subjected to human management.

Genetic structure analysis

The most likely number of homogeneous genetic clusters giving rise to the analyzed material was estimated with STRUCTURE. While peak levels for the associated likelihood value were observed between $K=3$ and $K=6$, the margin of error sharply increased beyond $K=4$. At the same time, the highest associated ΔK (the second-order rate of change of the likelihood function) was found for $K=3$ (482.23 vs. 197.46 for $K=2$ and 170.66 for $K=4$; insignificant for higher K values). These figures suggest that the material under study was shaped by three major genetic clusters (Fig. 2). In support of this hypothesis, the probability of assigning individuals to unique clusters was maximal for $K=3$, with 86.4 % of the individuals having a probability above 0.85 of belonging to one specific cluster. Only 4.1 % of the assignments showed probabilities under 0.65.

The above conclusions were bolstered by estimates of the fixation index (F_{ST}), a measure of differentiation caused by genetic structure (Weir and Cockerham 1984). Table 2 presents pairwise F_{ST} comparisons among the three STRUCTURE clusters, with higher values implying more differentiation. All tests were significant, with cluster 3 having the highest average value (0.434) and cluster 2 the lowest (0.244). The results of our F -statistic analysis were complemented by AMOVA data apportioning intra-cluster and inter-cluster genetic diversity. As shown in Table 3, three fourths of the overall variance were accounted for intra-cluster differences (Vb), whereas inter-cluster ones (Va) explained the rest. The overall F_{ST} value computed from the covariance components, 0.294, suggests a considerable degree of differentiation in El Bierzo.

To determine whether or not there was spatial structure to genetic variation, the correlation between genetic distance and geographic distance was analyzed. Whereas Mantel's test yielded a statistically significant value, the correlation coefficient between both distances was as low as 0.14 (Fig. 3).

Table 2 Pairwise F_{ST} -based genetic distances for the three gene pools inferred with STRUCTURE

	Cluster 1	Cluster 2	Cluster 3
Cluster 1	—		
Cluster 2	0.174**	—	
Cluster 3	0.552**	0.315**	—

** p value <0.01 based on 1000 permutations

Table 3 Analysis of molecular variance among the three gene pools inferred by STRUCTURE

Source of variation	Sum of squares	Variance components	Percentage of variation	F_{ST}	p value ^a
Inter-group	120.346	0.697 (Va)	29.36	0.294	<0.0001
Intra-group	454.142	1.677 (Vb)	70.64		
Total	574.488	2.374			

F_{ST} fixation index

^aBased on 10,000 permutations

BAPS, a Bayesian clustering software that treats allele frequencies and genetically diverged groups as random variables, found the optimal partitioning of the dataset at four clusters. When these results are visualized as a Voronoi tessellation graph (Fig. 4), where the areas surrounding each data point are closer to it than to any other neighboring point, no clear spatial structure can be inferred—in agreement with Mantel's tests. Our data do not support the occurrence of significant spatial structure in El Bierzo, a likely consequence of intensive human management.

Genetic characterization of local nuts

An accurate genetic characterization of local nuts is a prerequisite to optimize the conservation of valuable genotypes. It should also contribute to ensure the quality of marketed nuts and nut products. In fact, the growing importance of this resource for the local economy has been a major driver for our study. We analyzed DNA samples from 62 nuts belonging to all denominations commercialized as El Bierzo chestnuts. Seven markers were sufficient to identify 61 individual

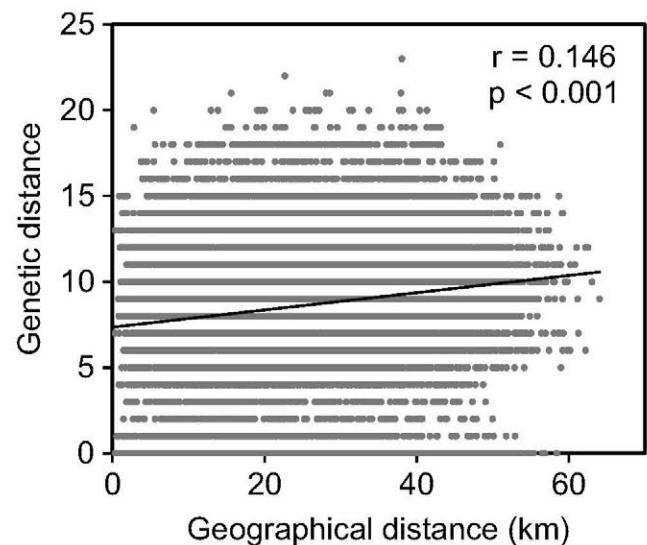
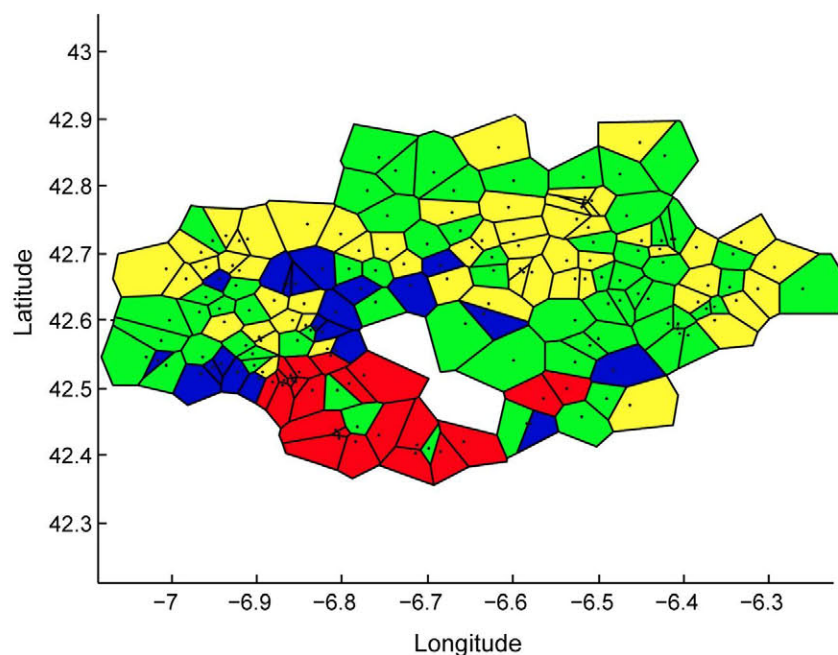


Fig. 3 Scatterplot of Nei's D_A distance vs. geographical distance for pairwise comparisons. Both variables were found to be only weakly correlated by Mantel's test (correlation coefficient 0.14)

Fig. 4 Spatial organization of the chestnut genotypes analyzed in this study. The diagram presents the Voronoi tessellation and assignment to the four clusters obtained with BAPS. Each color corresponds to a different cluster



genotypes (Table 4; see also Online Resources 3 and 4). The average values for H_e and H_o were 0.74 and 0.60, respectively, with CsCAT8 showing the largest difference between them. F_{IS} values suggested in this case an overall deficit of heterozygotes (average=0.13), with four loci showing significant deviations from Hardy-Weinberg equilibrium. Performing DNA amplification under less stringent conditions for a subset of homozygotes rendered essentially the same results.

To better define the genetic relationships among the pool of multilocus genotypes representing local denominations, a D_A distance matrix-based phylogram was constructed (Fig. 5a). Remarkably, its topology revealed three major groups. Cluster 1 was the largest, with 25 genotypes: Parede (15), Rapona or

Pelona (1), Mi Tierra (1), Injerta (4), Valdecañada (1), Verdella (1), Galeguiña (1), and Seroda (1). Cluster 2 included 24 genotypes: Negral (3), Roja (2), Courelá (3), Injerta (1), Raigona (3), Parede (2), Marela (3), De Presa (4), Valdecañada (1), Verdella (1), and Serodia (1). And cluster 3 included 13 genotypes: Galeguiña (2), Rapada (1), Rapiega (1), Marela Seroda (2), Parede (2), Mi Tierra (1), De Presa (1), Marela (1), Valdecañada (1), and Verdella (1). Several cases of misclassification by local producers were identified in this study, including Parede (4), De Presa (1), Marela (1), and Galeguiña (1). Nuts marketed as Verdella, Valdecañada, and Mi Tierra could not be assigned to any major group and did not even cluster together. To test the validity of these results, an independent clustering analysis was carried out using

Table 4 Summary statistics for the 6 SSR markers used to genotype 62 nuts from El Bierzo local varieties

Locus	Sample size	<i>A</i>	<i>R</i>	PIC	PI	<i>N</i>	<i>H_o</i>	<i>H_e</i>	<i>F_{IS}</i>
CsCAT2	62	8.90	4	0.71	9.5E-02	18	0.55	0.76	0.24**
CsCAT3	58	10.90	4	0.71	9.1E-03	18	0.60	0.74	0.17*
CsCAT6	57	7.99	4	0.75	7.2E-04	18	0.74	0.79	-0.04
CsCAT8	62	5.90	2	0.68	8.3E-05	13	0.50	0.74	0.22**
CsCAT17	60	6.93	2	0.74	7.7E-06	16	0.72	0.77	-0.05
EMCs38	56	9.00	4	0.61	1.1E-06	15	0.48	0.65	0.21**
Mean		8.27		0.70		16.33	0.60	0.74	0.13*

CsCAT and EMC loci were developed, respectively, by Marinoni et al. (2003) and Buck et al. (2003). Additional data are given in Online Resources 3 and 4

A allelic richness (based on a minimum sample size of 56 individuals), *R* number of rare alleles (<5 %), *PIC* polymorphism information content, *PI* probability of identity for increasing locus combinations, *N* number of different genotypes, *H_o* observed heterozygosity, *H_e* expected heterozygosity, *F_{IS}* Wright's inbreeding coefficient

p* value <0.05, *p* value <0.01 in locus-by-locus AMOVA for deviation from Hardy-Weinberg proportions

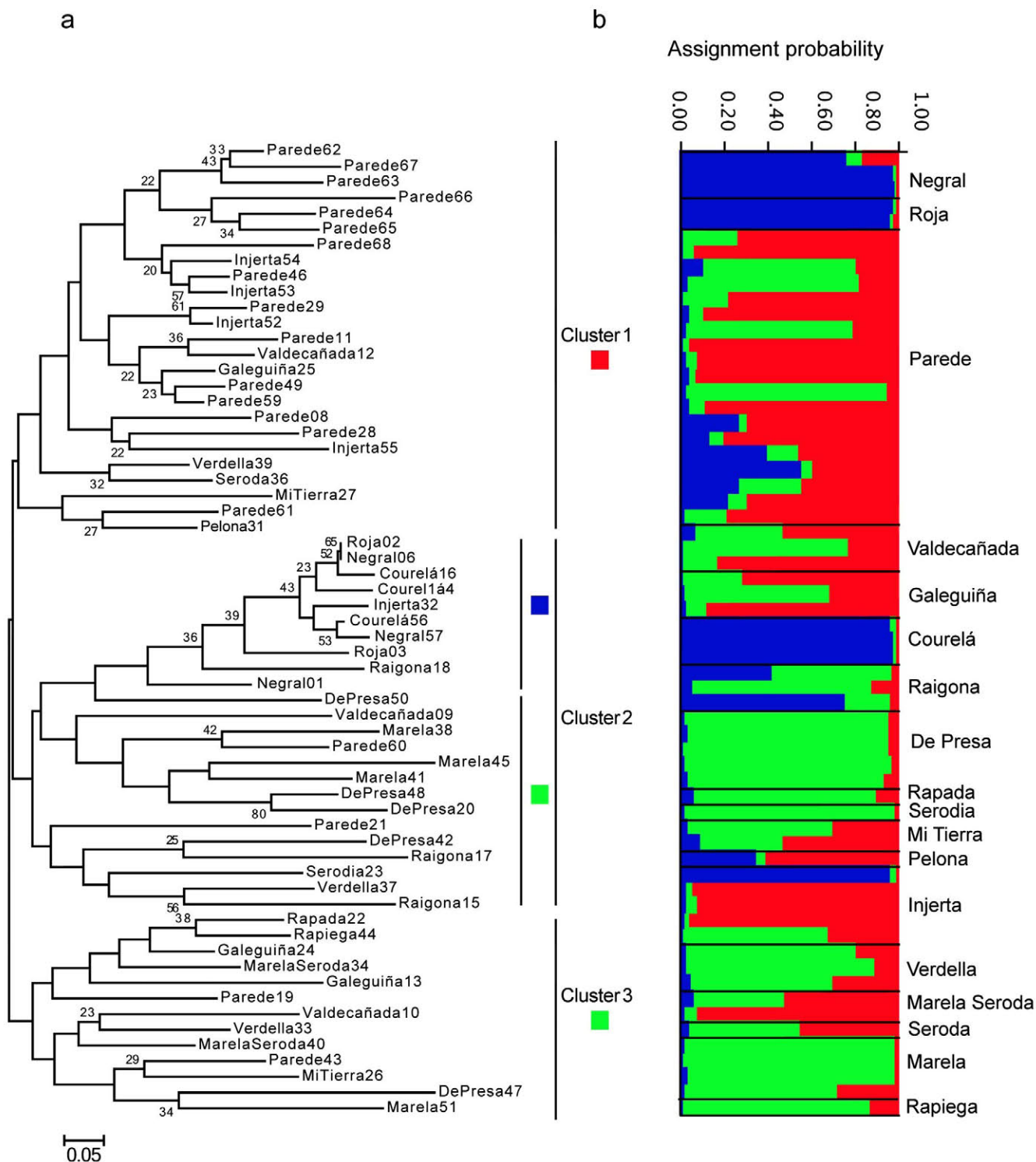


Fig. 5 Genetic relationships among 62 multilocus genotypes representing the 18 traditional nut denominations analyzed in this study. **a** Consensus D_A distance matrix-based phylogram obtained with the neighbor-joining algorithm (1000 replicates). *Bootstrap values* >20 %

STRUCTURE under a Bayesian approach (Fig. 5b). Three groups were inferred again, which overlapped to a large extent with those identified by the distance-based approach.

are indicated above the corresponding *branch*. **b** Quantitative analysis of genetic structure for the 62 genotypes, using the same method applied to leaf samples (see text and Fig. 2). Each *horizontal bar* represents a single individual analyzed

Table 5 shows within-group genetic variability calculations for the proposed clusters. The average number of alleles was 5.5 for the selected nuts. The number of alleles per group was quite similar, ranging from 4.83 to 6.17. Observed

Table 5 Within-group genetic variability at six loci for traditional nut cultivars (SE in parentheses)

Group	Traditional nomenclature	No. of accessions	Mean <i>N</i>	Ho	He	<i>F_{IS}</i>
Cluster 1	Injerta, Parede, Rapona or Pelona, Seroda	25	6.17 (1.83)	0.67 (0.15)	0.66 (0.06)	−0.04
Cluster 2	Courelá, Negral, Raigona, Roja, De Presa, Marela	24	5.50 (1.64)	0.65 (0.23)	0.71 (0.09)	0.47**
Cluster 3	Galeguiña, Rapada, Rapiega, Marela Seroda	13	4.83 (1.47)	0.55 (0.23)	0.62 (0.12)	0.22*

N number of alleles, *Ho* observed heterozygosity, *He* expected heterozygosity, *F_{IS}* inbreeding coefficient

***p* value <0.01 based on 10,000 permutations

heterozygosity was generally high, with an average value of 0.623. The highest value, 0.67, was found for cluster 1, followed by cluster 2 with 0.65 and cluster 3 with 0.55. Average expected heterozygosity ranged from 0.71 to 0.62, with cluster 2 showing the highest variability. *F_{IS}* values evidenced a significant deficit of heterozygotes in clusters 2 and 3 for the loci analyzed.

Discussion

We report here the first extensive genetic study of *C. sativa* in El Bierzo, a major nut production region in NW Spain with interesting features. The entire region is a tectonic depression enclosed by mountains that provide an unusual degree of genetic isolation and climatic conditions favorable for chestnut growth. Widely appreciated for their flavor and sweetness, El Bierzo nuts are officially protected. In this paper, selected microsatellite loci were used to analyze chestnut genetic structure and diversity, with an unprecedented sampling density (Fig. 1). Recent studies have analyzed grafted trees from this region, hinting at their relationships with other major production areas (Pereira-Lorenzo et al. 2010a, 2011; Martin et al. 2012). However, these works had a much broader geographical scope and did not include sufficient local samples for a detailed genetic characterization. We analyzed DNA from 169 grafted trees covering the entire region. In addition, we analyzed 62 nuts covering all local denominations. For both types of samples, a specific set of informative microsatellite markers was carefully selected among those developed by Marinoni et al. (2003) and Buck et al. (2003) (Tables 1 and 4). This set proved to be highly discriminant for the genotypes studied, giving combined PI values of 6.1×10^{-6} . When specific markers are considered, the discriminant capacity is higher than that reported by Marinoni et al. (2013) for chestnuts in Piedmont (Italy).

Genetic diversity in El Bierzo orchards

The loci employed in this study identified a total of 59 alleles (Table 1), with an overall richness comparable to that reported in studies covering much larger areas (Gobbin et al. 2007;

Wang et al. 2008; Martin et al. 2009; Mattioni et al. 2013; Lusini et al. 2014). Our marker set was highly informative for the material studied and a large proportion of alleles (61 %) was detected at a frequency below 5 %. Rare alleles are particularly useful for genotyping and also for establishing genome-wide associations (Dickson et al. 2010). In this respect, it must be noted that there are significant differences between El Bierzo and other European locations regarding the level of polymorphism of specific loci when comparable samples (c.g., grafted trees for nut production) are considered (Marinoni et al. 2003, 2013; Gobbin et al. 2007). From the resulting set of genotypes, a high degree of genetic diversity can be inferred for El Bierzo. Our data expand a recent study by Martin et al. (2012) reporting elevated levels of diversity among Spanish chestnut populations. This work included 15 trees from Las Medulas (western tip of El Bierzo), the largest study so far in this region. Our data also show that clonality due to grafting is about 42 %, a figure lower than previous estimates based on limited sampling (Pereira-Lorenzo et al. 2011). In addition to a relatively high heterozygosity (average *Ho* 0.61 and *He* 0.60), we found deviations from Hardy-Weinberg equilibrium for two loci. However, the average *F_{IS}* value (−0.06) suggests that heterozygosity loss, and hence long-term adaptability (Willi et al. 2006), is not a significant problem in El Bierzo.

Genetic structure and spatial autocorrelation

A clear understanding of local genetic structure is important to optimize both breeding and conservation strategies in major production sites. Unlike general reports on genetic structure across large geographic areas, this paper presents a fine-grained study focused on El Bierzo orchards. By following a Bayesian clustering approach, we concluded that the material under study derives from three major gene clusters. Calculations of *F_{ST}*-based genetic distances supported this hypothesis (Table 2) and suggested a considerable degree of differentiation among the three clusters. Moreover, unique pools were strong components of most genotypes under this scenario (Fig. 2). Several examples are known where crop genetic diversity is not abated by farming, but maintained in the form of traditional crop varieties or cultivars (Jarvis et al. 2008). Thus, the unexpectedly high genetic variability found

in El Bierzo may somehow reflect the effects of human intervention, irrespective of geographical isolation. Further analysis will be needed, however, to substantiate this notion. It is well established that both factors influence the genetic makeup of crops, especially when clonal effects take place (Jarvis et al. 2008; Pusadee et al. 2009; Myles et al. 2011). In the case of chestnut, domestication history has already been shown to affect genetic structure (Forneck 2005; Mattioni et al. 2008; Pereira-Lorenzo et al. 2010a).

It is noteworthy that most genotypes representing traditional nut varieties in El Bierzo could be readily assigned to one of the three aforementioned clusters (data not shown; discussed below). For grafted trees, a significant part of the observed genetic variation was explained by inter-cluster differences (Table 3). Like the F_{ST} comparisons discussed above, these data are supportive of an outstanding level of genetic differentiation, larger indeed than that estimated by Pereira-Lorenzo et al. (2010a) for grafted trees sampled across the entire Iberian Peninsula and Canary Islands. Two major origins have been postulated for cultivated chestnuts in the Iberian Peninsula: one in the north, to which all genotypes defined here would belong, and the other in the central zone (Pereira-Lorenzo et al. 2010a). The first pool is well differentiated from the rest of European production areas, as highlighted by Mattioni et al. (2008). Our genetic diversity data, summarized in Table 1, suggest that El Bierzo represents an important reservoir of variability within such pool.

Regarding spatial autocorrelation, two different strategies were applied in this paper (Figs. 3 and 4). Although a low positive correlation was found in Mantel's test, neither approach supported a solid relationship between genetic distance and geographic distribution. Most likely this reflects human intervention, which inevitably interacts with natural processes such as selection and drift to shape population dynamics. In line with these results, no significant geographic structure has been observed by Martin et al. (2012) upon analyzing 16 chestnut populations across Spain.

Genetic characterization of traditional nut varieties

As in other production areas, correct nut cataloguing is a major concern in El Bierzo. Local growers have relied so far on morphological and phenological traits, a system with well-known limitations (Bocacci et al. 2004; Pereira-Lorenzo et al. 2006a; and others). These are aggravated by hybridization, natural variability, and the lack of a core collection, resulting in incorrect cataloguing and frequent cases of homonymies and synonymies. This paper describes several examples where local denominations are inconsistent with genotyping. It also demonstrates the feasibility of developing reliable and cost-effective genotyping systems, based on selected microsatellite markers that are highly informative at the local level (Table 4). Extending it to local germplasm resources is

straightforward (compare Tables 1 and 4), an interesting target on its own (McCleary et al. 2013; Marinoni et al. 2013; Beghé et al. 2013).

By analyzing traditional chestnut varieties from Portugal and Spain, 10 panmitic populations were recently proposed by Pereira-Lorenzo et al. (2011). Four El Bierzo genotypes were assigned to three of those populations. Our D_A distance matrix-based phylogram for 18 local denominations (Fig. 5a) suggests the existence of three main clusters, a conclusion supported by Bayesian analysis of the same multilocus genotypes with STRUCTURE (Fig. 5b). The allelic richness of the three clusters is comparable (Table 5). Instances of misclassification by local producers were identified in all clusters. By using isoenzymes and morphological characters, Ramos-Cabrer and Pereira-Lorenzo (2005) and Pereira-Lorenzo et al. (2006a) had already proposed that Negral was in a different group than Injerta and Parede, as we show here. For these authors, Rapiega and Verdella were situated in the same group and Marela in a different one. This is in agreement with our classification. However, our data do not support the proposal that Galeguña clusters with Parede and Injerta. Such inconsistencies underscore the need to incorporate molecular markers to current classification systems, in El Bierzo and elsewhere. A set of well-defined genetic pools representing traditional varieties should be the target. As our paper shows, extensive local sampling can greatly contribute to this task, mainly in cultivation sites where genetic diversity is high. This would also help refine previous quantitative estimates (clonality, heterozygosity, and others) based on small sample sizes.

In conclusion, the results presented here evidence the need for more in-depth local genetic studies in Spain and elsewhere, particularly for major chestnut cultivation sites. They also put forth the need to complement traditional classification systems with molecular tools. Both steps will be necessary to improve current conservation and breeding strategies, as the recent study by Marinoni et al. (2013) in Piedmont so well exemplifies. They will also help understand how the practices of local growers and natural processes interact to shape genetic structure in places like El Bierzo, where diversity remains remarkably high after many decades of management.

Acknowledgments This research was financed by the Foundation *Ciudad de la Energía* (CIUDEN) from the Spanish Government. Julia Quintana was the recipient of a predoctoral fellowship from CIUDEN. We thank Dr. Pablo González-Jara for helpful comments and Begoña Prieto for excellent technical assistance. We are indebted to Drs. Avelino García and Luis del Riego for their advice and support throughout this work.

Data archiving statement The allelic profiles for chestnut varieties have been deposited in the TreeGenes database with accession number TGDR034.

References

- Bagnoli F, Vendramin GG, Buonamici A, Doulis AG, Gonzalez-Martinez SC, La Porta N, Magri D, Rads P, Sebastiani F, Fineschi S (2009) Is *Cupressus sempervirens* native in Italy? An answer from genetic and palaeobotanical data. *Mol Ecol* 18:2276–2286. doi:10.1111/j.1365-294X.2009.04182.x
- Barreneche T, Casasoli M, Russell K, Akkarak A, Meddour H, Plomion C, Villani F, Kremer A (2004) Comparative mapping between *Quercus* and *Castanea* using simple-sequence repeats (SSRs). *Theor Appl Genet* 108:558–566. doi:10.1007/s00122-003-1462-2
- Beghé D, Ganino T, Dall'Asta C, Silvanini A, Cirlini M, Fabbri A (2013) Identification and characterization of ancient Italian chestnut using nuclear microsatellite markers. *Sci Hortic* 164:50–57. doi:10.1016/j.scienta.2013.09.009
- Bocacci P, Akkarak A, Marinoni DT, Bounous G, Botta R (2004) Typing european chestnut (*Castanea sativa* Mill.) cultivars using oak simple sequence repeat markers. *Hortic Sci* 39:1212–1216.
- Brown AHD, Spillane C (1999) Implementing core collections. Principles, procedures, progress, problems, promise. In: Johnson RC, Hodgkin T (eds) Core collections for today and tomorrow. IPGRI, Rome, pp 1–9
- Buck EJ, Hadonou M, James CJ, Blakesley D, Russell K (2003) Isolation and characterization of polymorphic microsatellites in European chestnut (*Castanea sativa* Mill.). *Mol Ecol Notes* 3:239–241. doi:10.1046/j.1471-8286.2003.00410.x
- Casasoli M, Derory J, Morera-Dutrey C, Akkarak A, Meddour H, Plomion C, Villani F, Kremer A (2006) Comparison of quantitative trait loci for adaptive traits between oak and chestnut based on an expressed sequence tag consensus map. *Genetics* 172:533–546. doi:10.1534/genetics.105.048439
- Chakraborty R, De Andrade M, Daiger SP, Budowle B (1992) Apparent heterozygote deficiencies observed in DNA typing and their implications in forensic applications. *Ann Hum Genet* 56:455–457. doi:10.1111/j.1469-1809.1992.tb01128.x
- Cheng L, Connor TR, Sirén J, Aanensen DM, Corander J (2013) Hierarchical and spatially explicit clustering of DNA sequences with BAPS software. *Mol Biol Evol* 30:1224–1228. doi:10.1093/molbev/mst028
- Conedera M, Krebs P, Tinner W, Pradella M, Torriani D (2004) The cultivation of *Castanea sativa* (Mill.) in Europe, from its origin to its diffusion on a continental scale. *Veget Hist Archaeobot* 13:161–179. doi:10.1007/s00334-004-0038-7
- Corander J, Sirén J, Arjas E (2008) Bayesian spatial modeling of genetic population structure. *Comput Stat* 23:111–129. doi:10.1007/s00180-007-0072-x
- Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB (2010) Rare variants create synthetic genome-wide associations. *PLoS Biol* 8:e1000294. doi:10.1371/journal.pbio.1000294
- Earl D, von Holdt B (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4:359–361. doi:10.1007/s12686-011-9548-7
- Ellegren H (2004) Microsatellites: simple sequences with complex evolution. *Nat Rev Genet* 5:435–445. doi:10.1038/nrg1348
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620
- Excoffier L, Lischer HEL (2010) Arlequin suite version 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 10:564–567. doi:10.1111/j.1755-0998.2010.02847.x
- Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783–791. doi:10.2307/2408678
- Forneck A (2005) Plant breeding: clonality. A concept for stability and variability during vegetative propagation. In: Esser K, Lüttge U, Beyschlag W, Murata J (eds) Progress in botany, vol 66. Springer, Berlin Heidelberg, pp 164–183
- Gobbin D, Hohl L, Conza L, Jermimi M, Gessler C, Conedera M (2007) Microsatellite-based characterization of the *Castanea sativa* cultivar heritage of southern Switzerland. *Genome* 50:1089–1103. doi:10.1139/G07-086
- Goudet J (1995) FSTAT: a computer program to calculate F-statistics. *J Hered* 86:485–486.
- Guo X, Elston RC (1999) Linkage information content of polymorphic genetic markers. *Hum Hered* 49:112–118. doi:10.1159/000022855
- Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Mol Ecol Resour* 9:1322–1332. doi:10.1111/j.1755-0998.2009.02591.x
- Jarvis DI, Brown AHD, Cuong PH, Collado-Panduro L et al (2008) A global perspective of the richness and evenness of traditional crop-variety diversity maintained by farming communities. *Proc Natl Acad Sci U S A* 105:5326–5331. doi:10.1073/pnas.0800607105
- Krebs P, Conedera M, Pradella M, Torriani D, Felber M, Tinner W (2004) Quaternary refugia of the sweet chestnut (*Castanea sativa* Mill.): an extended palynological approach. *Veget Hist Archaeobot* 13:145–160. doi:10.1007/s00334-004-0041-z
- Kremer A, Abbott A, Carlson J, Manos PS, Plomion C, Sisco P, Staton ME, Ueno S, Vendramin GG (2012) Genomics of Fagaceae. *Tree Genet Genomes* 8:583–610. doi:10.1007/s11295-012-0498-3
- Lang P, Dane F, Kubisiak TL, Huang H (2007) Molecular evidence for an Asian origin and a unique westward migration of species in the genus *Castanea* via Europe to North America. *Mol Phylogenet Evol* 43:49–59. doi:10.1016/j.ympev.2006.07.022
- Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 21:2128–2129. doi:10.1093/bioinformatics/bti282
- Lusini I, Velichkov I, Pollegioni P, Chiochini F, Hinkov G, Zlatanov T, Cherubini M, Mattioni C (2014) Estimating the genetic diversity and spatial structure of Bulgarian *Castanea sativa* populations by SSRs: implications for conservation. *Conserv Genet* 15:283–293. doi:10.1007/s10592-013-0537-0
- Magri D, Fineschi S, Bellarosa R, Buonamici A, Sebastiani F, Schirone B, Simeone MC, Vendramin GG (2007) The distribution of *Quercus* suber chloroplast haplotypes matches the palaeogeographical history of the western Mediterranean. *Mol Ecol* 16:5259–5266. doi:10.1111/j.1365-294X.2007.03587.x
- Marinoni D, Akkarak A, Bounous G, Edwards KJ, Botta R (2003) Development and characterization of microsatellite markers in *Castanea sativa* (Mill.). *Mol Breed* 11:127–136. doi:10.1023/a:1022456013692
- Marinoni D, Akkarak A, Beltramo C et al (2013) Genetic and morphological characterization of chestnut (*Castanea sativa* Mill.) germplasm in Piedmont (north-western Italy). *Tree Genet Genomes* 9:1017–1030. doi:10.1007/s11295-013-0613-0
- Martin MA, Alvarez JB, Mattioni C, Cherubini M, Villani F, Martin LM (2009) Identification and characterisation of traditional chestnut varieties of southern Spain using morphological and simple sequence repeat (SSRs) markers. *Ann Appl Biol* 154:389–398. doi:10.1111/j.1744-7348.2008.00309.x
- Martin MA, Mattioni C, Cherubini M, Turchini D, Villani F (2010) Genetic diversity in European chestnut populations by means of genomic and genic microsatellite markers. *Tree Genet Genomes* 6:735–744. doi:10.1007/s11295-010-0287-9
- Martin MA, Mattioni C, Molina J, Alvarez J, Cherubini M, Herrera M, Villani F, Martin L (2012) Landscape genetic structure of chestnut (*Castanea sativa* Mill.) in Spain. *Tree Genet Genomes* 8:127–136. doi:10.1007/s11295-011-0427-x
- Mattioni C, Cherubini M, Micheli E, Villani F, Bucci G (2008) Role of domestication in shaping *Castanea sativa* genetic variation in Europe. *Tree Genet Genomes* 4:563–574. doi:10.1007/s11295-008-0132-6

- Mattioni C, Martin MA, Pollegioni P, Cherubini M, Villani F (2013) Microsatellite markers reveal a strong geographical structure in European populations of *Castanea sativa* (Fagaceae): evidence for multiple glacial refugia. *Am J Bot* 100:951–961. doi:10.3732/ajb.1200194
- McCleary T, McAllister M, Coggeshall M, Romero-Severson J (2013) EST-SSR markers reveal synonymies, homonymies and relationships inconsistent with putative pedigrees in chestnut cultivars. *Genet Resour Crop Evol* 60:1209–1222. doi:10.1007/s10722-012-9912-9
- Myles S, Boyko AR, Owens CL et al (2011) Genetic structure and domestication history of the grape. *Proc Natl Acad Sci U S A* 108:3530–3535. doi:10.1073/pnas.1009363108
- Neale DB, Kremer A (2011) Forest tree genomics: growing resources and applications. *Nat Rev Genet* 12:111–122. doi:10.1038/nrg2931
- Nei M, Tajima F, Tatenio Y (1983) Accuracy of estimated phylogenetic trees from molecular data. *J Mol Evol* 19:153–170. doi:10.1007/bf02300753
- Peakall R, Smouse PE (2006) GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6:288–295. doi:10.1111/j.1471-8286.2005.01155.x
- Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28:2537–2539. doi:10.1093/bioinformatics/bts460
- Pereira-Lorenzo S, Díaz-Hernández B, Ramos-Cabrer AM (2006a) Use of highly discriminating morphological characters and isozymes in the study of Spanish chestnut cultivars. *J Am Soc Hortic Sci* 131:770–779. doi:10.1311/6/770.full
- Pereira-Lorenzo S, Ramos-Cabrer AM, Díaz-Hernández MB, Ciordia-Ara M, Ríos-Mesa D (2006b) Chemical composition of chestnut cultivars from Spain. *Sci Hortic* 107:306–314. doi:10.1016/j.scienta.2005.08.008
- Pereira-Lorenzo S, Costa R, Ramos-Cabrer AM, Ribeiro C, da Silva M, Manzano G, Barreneche T (2010a) Variation in grafted European chestnut and hybrids by microsatellites reveals two main origins in the Iberian Peninsula. *Tree Genet Genomes* 6:701–715. doi:10.1007/s11295-010-0285-y
- Pereira-Lorenzo S, Díaz-Hernández B, Ramos-Cabrer A (2010b) Chestnut in Spain, from nut and timber production to industry. In: Bounous G (ed) *I European Congress on chestnut—Castanea*. ISHS Acta Horticulturae, Cuneo–Torino, pp 499–503
- Pereira-Lorenzo S, Costa R, Ramos-Cabrer AM, Ciordia-Ara M, Ribeiro CAM, Borges O, Barreneche T (2011) Chestnut cultivar diversification process in the Iberian Peninsula, Canary Islands, and Azores. *Genome* 54:301–315. doi:10.1139/g10-122
- Petit RJ, Carlson J, Curtu AL, Loustau ML, Plomion C, González-Rodríguez A, Sork V, Ducousso A (2013) Fagaceae trees as models to integrate ecology, evolution and genomics. *New Phytol* 197:369–371. doi:10.1111/nph.12089
- Pollegioni P, Woeste K, Olimpieri I, Marandola D, Cannata F, Emilia Malvolti M (2011) Long-term human impacts on genetic structure of Italian walnut inferred by SSR markers. *Tree Genet Genomes* 7:707–723. doi:10.1007/s11295-011-0368-4
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Pusadee T, Jamjod S, Chiang Y-C, Rerkasem B, Schaal BA (2009) Genetic structure and isolation by distance in a landrace of Thai rice. *Proc Natl Acad Sci U S A* 106:13880–13885. doi:10.1073/pnas.0906720106
- Ramos-Cabrer AM, Pereira-Lorenzo S (2005) Genetic relationship between *Castanea sativa* Mill. trees from northwestern to south Spain based on morphological traits and isoenzymes. *Genet Resour Crop Evol* 52:879–890. doi:10.1007/s10722-003-6094-5
- Rubio A (2009) *Castanea sativa* forests. In: Spain Government—MAGRAMA (ed) *Ecological basis to preserve habitat types of EU interest in Spain*. Ministerio de Agricultura, Alimentación y Medio Ambiente, Madrid, pp 1–64
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425
- Takezaki N, Nei M, Tamura K (2010) POPTREE2: software for constructing population trees from allele frequency data and computing other population statistics with Windows interface. *Mol Biol Evol* 27:747–752. doi:10.1093/molbev/msp312
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28:2731–2739. doi:10.1093/molbev/msr121
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* 4:535–538. doi:10.1111/j.1471-8286.2004.00684.x
- Wang Y, Kang M, Huang H (2008) Microsatellite loci transferability in chestnut. *J Am Soc Hortic Sci* 133:692–700
- Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38:1358–1370. doi:10.2307/2408641
- Willi Y, Van Buskirk J, Hoffmann AA (2006) Limits to the adaptive potential of small populations. *Annu Rev Ecol Syst* 37:433–458. doi:10.1146/annurev.ecolsys.37.091305.110145